Report of the 29th Met Office Scientific Advisory Committee Meeting (21-23rd January 2025)

Annex: Recommendations and Questions

Science and Technology Overview

The Met Office is currently undergoing a transformative phase, marked by numerous external and internal changes. New supercomputer capabilities (Gen1a, b, and eventually Gen2) are being delivered and deployed. Successfully porting the codes for these systems is a monumental effort, and we congratulate the Met Office staff on this achievement. This progress enables the implementation of strategic initiatives, such as releasing parallel suites and advancing climate modeling strategies.

Simultaneously, the disruption brought about by AI and machine learning (AI/ML) is taking hold in the meteorological and climate communities. In response, the Met Office has deprioritized several activities, freeing up resources to engage more robustly in AI/ML. This decision by the Met Office's leadership is commendable, though it is important to acknowledge the impact it has on staff and community contributions. The Met Office is advised to maintain strong connections with partners to ensure that both UK and international capabilities remain intact.

The Met Office's efforts in applying AI/ML are already yielding results, as seen in the AI4NWP and AI4CLIMATE programs. However, it is evident that the Met Office, like many institutions worldwide, is still exploring various options. Therefore, a more delineated AI/ML strategy specific to the new strategy period could be beneficial. While the MOSAC remains neutral to the choices made, it suggests developing a roadmap with critical points highlighting the fields in which the Met Office aims to be a world leader, a fast follower, or a follower, and what this implies for resources and partnerships.

The MOSAC does not view this as a binary decision and encourages leveraging the excellent science in physics at the Met Office. Currently, the Met Office is developing FastNet to stay competitive, but it could also benefit from ECMWF's efforts. Given the Met Office's world-leading expertise in physics, hybrid physics-based and data-driven approaches may be a good fit. Likely, AI/ML will play a significant role in downstream activities towards application. The choices made in the strategy will also influence the selection of partnerships.

The Met Office is encouraged to develop strategic personnel planning, aligning resource decisions and staff competencies with the aforementioned strategy.

STO-R1 MOSAC recommends that the Met Office present a comprehensive AI/ML strategy, complete with a roadmap and associated resource planning reflection. This strategy could articulate areas where the Met Office aims to be a world leader, a fast follower, or an applier of AI/ML.

We welcome this recommendation and agree that an AI/ML roadmap will be timely. This roadmap will identify activities that track a course towards the Met Office's stated ambitions of:

1. Maintain the UK's position as a world-leading modelling centre by remaining at the cutting edge of both physics-based and ML-based modelling.
2. Deliver world-leading weather and climate prediction data through an optimal blend of physics-based and ML-based modelling.
3. Position the Met Office as a centre of excellence by stimulating the national and international communities into improving the accuracy of, and trust in, ML weather models by establishing the criteria for design of models and frameworks for assessment and assurance of their quality.

MOSAC recognizes that data management is both complex and crucial for the mission and strategy of the Met Office. MOSAC encourages the Met Office to enhance coherence in its data management strategy and implementation, considering AI/ML developments to fully exploit its unique position with domain-specific databases.

Additionally, MOSAC acknowledges that data infrastructure and management are essential for the Met Office's mission and strategy across all activities. While computational advancements with Gen 1 and Gen 2 are clear, the same cannot be said for data.

STO-R2 MOSAC recommends that the Met Office strengthen coherence in its data infrastructure and data management strategy, considering AI/ML developments for domain-specific databases.

We thank MOSAC for these recommendations. This will be a focus of the newly formed Data Curation team, alongside Data Engineering. Work in this area is also ongoing with the Data Delivery programme Data Maturity project, where data governance and metadata is a priority, and in the delivery and future improvements plan for Azure MASS which will enable improved data governance (e.g. through improved metadata and cataloguing).

MOSAC supports the approach to provide more staffing resource for AI developments. However, we would like a bit more information on the strategy for filling the gaps left in research observations, which are important both for model development (e.g., focussed field campaigns) and in providing proof of concept/evidence for investments in new operational observations. We agree that it may be possible to outsource observing campaigns from universities and other research partners. However, there is a risk that the current financial climate for this sector in the UK is likely to result in some reduction of investment in new instruments, loss of expertise and other instabilities.

STO- Q1 How will the Met Office ensure that any commissioning is realistically costed, and an attractive opportunity for research partners to bid for?

We agree that research observations will continue to be important for model development. Our future approach in this area is to be research observation users, rather than makers, and this will entail enhancing our partnerships. One aspect of this will involve commissioning work. The Met Office has well established procurement processes for commissioning, including for research work. The process involves “early market engagement” where we share information in advance and collate feedback from interested parties. This feedback helps to shape the opportunities, including the scope and cost. We will consider feedback from the community as we continue to evolve our approach to collaboration.

STO-R3 More generally, MOSAC recommends that the Met Office map its dependencies on research partners and use this evidence to lobby UKRI and other funders to support these areas through national capability funding (e.g., at NERC centres) or directed programmes.

We thank MOSAC for these suggestions and will be considering dependencies on external partners. We agree that there will be a need to form a national conversation on UK capabilities in weather and climate

Supercomputer Data Exploitation

While the market is uncertain, the growth in applying AI/ML by the Met Office is a certainty. Various papers and presentations made clear that AI/ML will, regardless of the precise strategy,

be part of Met Office's workflow from foundational science to operations and services. Hence GPU resources in Met Office's computing capabilities are a likely requirement. The MOSAC would like to emphasize that sufficient capabilities need to be present before and as part of Gen 2 to keep progressing these activities and ensuring the quality.

SDE-R1 MOSAC recommends keeping developing software infrastructures, using modern software development paradigms, that allow for flexible hardware infrastructures and integration of data management and computation.

We agree that whilst many factors in this area are uncertain, one thing we do know is that there is a growing need for flexibility. Our approach to procuring our compute for both our scientific and product generation/dissemination is increasingly based on engagement with the private sector and adopting industry standards, whilst recognising where there are specialist use-cases specific to our sector and requirements. We agree that this means adapting our approaches, which the Met Office Science and Technology teams are working on together.

SDE-R2 MOSAC would welcome, if possible, specs on the computing infrastructures (Gen1 and 2), including scaling graphs and other performance metrics (for example as annex to the material presented).

We will be happy to provide specifications for Generation 1 systems with the papers for next year. We will aim to include performance comparisons of some of the key models between old and new systems which we believe to be a more relevant way of considering performance than abstract benchmarks such as HPL or HPCG. We will unfortunately not be able to provide specifications for Generation 2 as we don't expect to have them in any detail by that point.

Data is key to meteorology and climate R&D and services, and this is even more apparent in the transformation towards more data-driven science and downstream services.

SDE-R3 The MOSAC recognizes that a data strategy is needed and implemented to improve the overall data management at the Met Office. In various papers and presentations data was central, but the data strategy shows less coherence. When considering Met Office science, particular attention is needed for the research needs with partners (upstream).

We will be working on improving data management within research, following our Met Office Data Strategy Framework high level principles and data readiness categories, Data Delivery programme metadata and governance, and security and IPR considerations. We expect that AI/ML approaches and opportunities from Azure MASS and Git Migration will act as good vehicles to improve data management generally in 2025. We will be happy to report on progress in future meetings.

AI for National Capability

The committee is pleased to see that the reprioritization has resulted in more resources being allocated to machine learning work and that the previous delays concerning the Turing partnership have been resolved. This development provides grounds for optimism that progress can now accelerate.

However, MOSAC would appreciate further clarification regarding the articulation of the AI/ML strategy across the Met Office. While there are notable examples of innovative work, such as

post-processing and physics emulation, how these initiatives fit into the broader strategy and their impact on observation and data curation remain unclear.
For instance, the committee found the detailed plans for FastNet somewhat ambiguous beyond the reasonable goals of improving sharpness through loss function adjustments, developing ensembles, and enhancing high-resolution capabilities. Clarification on the realistic progress expected over the next six to twelve months, especially compared to advancements made by big tech companies or other entities in the meteorological community, would be valuable. Additionally, more clarity on the aspirations beyond atmospheric studies, such as land, ocean, and wave modeling, would be welcomed. Also, with global resolutions increasing and AI/ML demonstrating strong potential for refinement at local scales, how does the Met Office envision the relevance and future of regional systems?

More generally, the committee thought that the Met Office could usefully give more thought to its positioning in this fast moving landscape, scoping out in more detail in what areas it wishes to truly lead, where it aspires to be a 'fast follower' (and what that means in more detail) and where it is happy to be more reliant on the efforts of others.

AI-R1 Hence MOSAC commends the effort so far but would welcome a paper on the AI/ML strategy next year and a more detailed plan for what they aspire to achieve in the next 6-12 months, and where this will position them with respect to others and their partnerships (e.g., Turing Institute, Momentum and UK traditional partners).

There was considerable discussion within the MOSAC about the challenges faced by the whole community in transitioning from impressive capabilities to their actual use in some services. Stepping more strongly into this space, through operational meteorologist assessments of models from a user perspective and the development of trailblazing products and service, feels like it may provide one of the best opportunities for the Met Office to play a leadership role.

The Met Office is currently updating its overarching strategy and during 2025 will refresh the Research and Innovation Strategy, which will incorporate AI/ML ambitions. In parallel we are developing an AI/ML roadmap that identifies activities tracking a course towards the Met Office's stated ambitions of:

1. Maintain the UK's position as a world-leading modelling centre by remaining at the cutting edge of both physics-based and ML-based modelling.
2. Deliver world-leading weather and climate prediction data through an optimal blend of physics-based and ML-based modelling.
3. Position the Met Office as a centre of excellence by stimulating the national and international communities into improving the accuracy of, and trust in, ML weather models by establishing the criteria for design of models and frameworks for assessment and assurance of their quality.

We shall work with both MOAP and Momentum partners to develop and deliver these ambitions. We would be happy to provide an update at a future MOSAC meeting.

AI-R2 MOSAC recommends acceleration of efforts to evaluate and exploit ML models (e.g., by operational meteorologists and trailblazing products).

We thank MOSAC for its guidance and will consider our approach to the evaluation of ML models, noting that resource remains constrained and that we need to prioritise carefully.

Seamless Extremes

MOSAC commends the Met Office's international leadership in the effective use of ensembles in forecasting operations, as demonstrated by the recently completed Ensemble Exploitation initiative. The new Ensemble Sensitivity Analysis (ESA) and Clustering tools provide valuable insights to forecasters, and the Met Office is encouraged to integrate these tools into their operational processes.

SE-R1: Operationalise the new ESA and clustering tools for operational forecasting and science investigation.

The Science and Services directorates are currently discussing possible pathways for operationalisation of the new ESA tools, recognising that these tools are still under trial with our operational users, and may require further science investigation following feedback. We have plans in the next financial year to continue exploration and improvement of the clustering tools.

SE-R2: Support learning, tools, and methods across different timescales and domains through fostering communities of practice, training, seminars, and temporary assignments.

We agree with this recommendation. We intend to continue with existing activities within the Met Office which help to address this, including the communities of practice, weather-climate secondments (e.g. linked to the Resilience and Preparedness to tropical cyclones across South Africa (REPRESA) project) and weather and climate scientist involvement in multi-year, cross-timescale research projects (e.g. the Horizon Europe Human-Tech Nexus Project).

Through the Horizon Europe project ASPECT the Met Office has been developing methodologies to produce seamless climate information across timescales, blending seasonal and decadal forecasts with climate projections. Case studies with the humanitarian and pensions sector in ASPECT have been building on these approaches to demonstrate its value to end users and enhance uptake of seamless sector specific spatial risk information. An associated spatial risk Community of Practice has been established with MOAP academics to exchange best practise and provide training in shared tools for assessment of spatial risk.

The relevance of understanding extremes and providing information to stakeholders to inform policies is clear and aligns with the strategic goals of Met Office. Examples of extreme events were presented across temporal and spatial scales, showcasing the need for a wide range of perspectives from different scientific domains and stakeholders. For instance, the Valencia flood event depended not only on rainfall intensity and duration but also on land use and the ability of authorities to inform the public and manage disaster responses. Despite the challenges, MOSAC recognizes the impressive results and invites the team to reflect on the risks of lasting changes that could shift risks and hazards.

SE-R3: The Met Office is encouraged to continue developing workflows that can be applied across various extreme events, leveraging its data and modeling capabilities and partnering with institutions that have expertise in impact data (such as damage curves, societal data, and local knowledge) to enhance the effectiveness of risk management tools and strategies.

Thank you for this encouragement. We intend to continue existing activities and partnerships to support this recommendation and will look to identify further opportunities moving forward. The Met Office has continued to enhance its use of CLIMADA (and event based spatial risk tool) including working with MOAP (UCL) to incorporate direct building simulations of 20,000 English

schools to enable a climate risk assessment of overheating, ultimately to inform DfE's climate change adaptation strategy. Similar assessments have also been produced for MoJ and exploratory work on overheating sub-stations for National Grid. Further development of this framework is planed through the UKRI Maximising Adaptation to Climate Change (MACC) programme, with focus on adaptation options assessment.

Furthermore, the National Institute of Health Research (NIHR) has just renewed the Met Office's involvement with a Health Protection Research Unit (HPRU) for another 5 years, from April 2025, working with LSHTM, UKHSA and other consortium partners. Alongside the NERC project THERM-UK this will be building the quantities evidence base for health impact of extreme events to inform risk management decisions.

Physics Development

The Met Office is a world leader in understanding physical processes in the weather and climate system, and in translating this understanding into parameterizations in numerical models. MOSAC is pleased to recognize that ssignificant progress has been made towards the unified moist physics configuration for GAL and RAL, with CoMorph-B significantly improving convection over the tropics, the CASIM microphysics scheme initialised with CAMS aerosols, and a single cloud scheme. This has been achieved in the particularly challenging environment of developing a next-generation model and porting it to a new supercomputer. MOSAC has been critical of the inability to bring science advances into the operational system during this period. With this background, we are pleased to see that there is now a plausible timetable for the introduction of improved physics in PS47 and PS48 this year and next.

Looking forward, MOSAC supports the continuation of physics development, which consists of incorporating ML while consolidating traditional developments to achieve an optimal hybrid physical-AI blend to achieve the most accurate representation of the atmosphere and land surface. A path has been laid out for future developments, including the scale-aware version of CoMorph, and the unification of closely related schemes, and we look forward to seeing the results of this work.

PD-R1: Given the continuing importance of the physically based model alongside new AI methods, MOSAC recommends that sufficient resources be retained for traditional parameterization development to sustain this work.

We recognise the continued importance of physics based modelling and thank MOSAC for this recommendation. We are committed to continued effort on physical parametrizations.

PD-R2: Given the excellent progress already made toward unification between global and regional physics, MOSAC recommends that the important ongoing work on hybrid AI-physical modelling should not impede the physics unification efforts.

We thank MOSAC for their supportive comments and intend to prioritise and resource in this way.

Equally important is the progress made in hybrid AI-physical modelling, by combining traditional model development with emulated elements, and exploring a wide range of original activities, based on a framework for embedding machine-learnt algorithms into physical models. Some of these developments exploit the unified modelling approach to emulate kilometer-scale simulations and feed the result into the climate model. MOSAC commends the Met Office on the

very innovative attempts to explore a range of hybrid physical-AI modelling approaches for some parts of the physics package with the aim of improving upon the existing physics and combining traditional model development with emulated elements. A number of technical hurdles have been overcome by embedding machine-learnt algorithms into the physical models, paving the way for future advances.

PD-R3: As there is no mention of work on the surface schemes, MOSAC recommends considering the emulation of the surface or one of the surface schemes, to assess the feasibility of this approach.

There is research, both at the Met Office and amongst partners, emulating parts of the JULES land surface model. Reporting on them was omitted this year due to space considerations, but a plenary presentation on land surface progress and plans may be appropriate next year.

The challenges of recent years have also highlighted the broader problem of the long pathway from improved understanding to operational deployment, and the many vetoes points along the way. MOSAC is pleased to see that this problem is being addressed. In the short term, the use of Trailblazers will allow at least some users to get the benefit of recent advances without waiting for all possible impacts to be tested. The plans to allow incremental updates to the model on the Gen2 supercomputer provide a potential alternative model to the sometimes-cumbersome Parallel Suites.

PD-R4: We encourage the use of an agile approach to introduce cutting-edge science to operational systems where appropriate.

We agree, and with the pace of AI developments in particular, we recognise the time from development to operational implementation needs to reduce. This will require striking the right balance between ensuring our model, and the products and services they feed, are robust, whilst not being excessively risk averse (recognising that even at the moment we can't eliminate all risk). There is also still a need to prioritise NGMS implementation so more significant work to improve agility (beyond the trailblazer implementation) is likely to be post-PS49.

Observation Exploitation

MOSAC was delighted to see that the successful implementation of JOPA is now enabling the pull through of new observations into operations, and that JOPA is enabling the acceleration of this pull-through. We particularly congratulate the Met Office on their successes with Marine Gliders and global Mode-S, a success highlighted at one of the presentations and are pleased to see that there are plans to bring in some key new satellite data over the coming years.

MOSAC believes that it is extremely important for the Met Office to work on upgrading observations databases and other data infrastructure so that they are fit for the ingestion of heterogeneous datatypes into operational NWP and nowcasting, including high volume, high velocity data expected from future satellite missions, crowdsourcing etc. Such work should include a data management strategy across Met Office activities. MOSAC was pleased to see that there is work planned on a strategy for this already.

OE-R1 MOSAC would like to see an end-to-end observations data infrastructure and data management strategy, so that the strategy for operational receipt and ingestion of data is joined

up with the new data initiatives for downstream services (paper 29.02) and input datasets work for AI/ML (paper 29.3). Such a strategy can be extended to data management strategy for model data and services products.

The “Future Observations Data Systems Roadmap" project (FODSR) has been set up to define a future architecture for observations data that is designed to meet the future requirements of all observations use across the office including operational meteorology, weather prediction (physical and AI), marine modelling, climate, post processing, nowcasting and data services. This project will align with existing Data Delivery projects that manage the architecture for data management for exploitation downstream of the modelling environment and inform the new AI/ML focussed efforts mentioned in response to STO-R2. We anticipate having clarity on direction by the end of 2025.

MOSAC was asked to give a view on how the Met Office can work better with MOAP and UKRI to enable academic research that facilitates faster pull-through of new observations to operations. MOSAC notes that observations technology development and data assimilation topics are difficult to find funding for in UKRI “responsive mode”, where there are cultures and expectations around the nature of suitable proposals, but there are recent examples where the Met Office has successfully influenced calls for directed programmes. Similar strategies have also been employed overseas (e.g., NOAA’s data assimilation research consortium). Secondly, while in principle the new JEDI software should make collaboration and pull-through easier, it is difficult to install in university computer systems, and, once installed, not necessarily straightforward for early career researchers to pick up and use. The Met Office may need to support porting of JEDI, together with appropriate data to external computer systems to enable this type of research.

OE-R2 MOSAC recommends that the Met Office use its influence and partner with UKRI in creation of directed funding programmes.

We thank MOSAC for this helpful insight. The Met Office is currently engaged with Reading University to port and install JEDI as part a Data Assimilation training course, to be run every year. This is the first step in JEDI’s full deployment, enabling university staff and students to perform their research within the JEDI framework, so enabling faster pull through from Research to Operations and developing a talent-pipeline of experienced users. The experience gained here will allow JEDI to be more easily installed elsewhere in future, and will help as we engage with UKRI on potential future co-developed research programmes.

Forecast Accuracy

MOSAC commends the Met Office for looking strategically at its verification and evaluation approaches for measuring the accuracy of models, both numerical and AI/ML. Assessing forecast accuracy is necessary for understanding performance, identifying where improvements must be made, and building user trust in the forecast. MOSAC endorses the MO’s intention to include additional user value-oriented metrics in its evaluation processes. When a new model comes along, for example, this information can be very useful to share with users.

FA-R1: Seek opportunities to measure the quality of forecasts (and forecast improvements) in terms of value to the user. This would likely involve subject matter experts in the Programmes and Services directorates.

We thank MOSAC for these supportive comments and we are pleased our aspirations to bring more user-value metrics into the development process have been well received. We are increasingly discussing our approach to verification and evaluation across the "value chain" and joining up subject matter experts on this topic.

MOSAC congratulates the Met Office on this year's release of the IMPROVER post-processing system which underpins public weather forecasts on the web and app. It was pleasing to see that the use of neighbourhood-based methods to correct model biases led to more accurate forecasts (both actual and perceived) from IMPROVER as compared to Best Data. MOSAC supports the planned use of the SAMOS method for ensemble bias correction to further improve IMPROVER accuracy.

MOSAC is comfortable with the current approach to implement model upgrades even when it may lead to slight degradation in the NWP score. It is understood that the advantages of improved physical realism, reliability, seamlessness, and technical advances in model structure may outweigh the disadvantage of slightly less accurate forecasts. The existing model acceptance criteria appear to cover this situation.

To speed the pull through of innovation to services, the Met Office could consider releasing a "beta" version of a model/system to internal and trusted external users for their own evaluation. This could potentially involve a gateway process that demonstrates that the system meets certain accuracy and reliability criteria (which could be more relaxed than the final acceptance criteria), with the expectation that further testing and possible refinement would still be needed before the final release. This contrasts with the Trailblazer releases which are more in the research realm.

The rapid rise of AI/ML-based weather and climate models has created the need to verify and evaluate such models for their accuracy, realism and utility. MOSAC endorses the proposed 3-pronged approach to apply

- similar verification approaches as currently used for numerical models (augmented by metrics for variability and consistency);
- context-specific approaches that consider how the forecasts will be used further down the value chain (for example, to inform hazard forecasts and warnings), and
- forecaster evaluation that applies expert knowledge to assess the realism of the multi-variate forecast.

FA-R2: The Met Office should actively develop and demonstrate a comprehensive approach for evaluating the accuracy and value of AI/ML models, in cooperation with relevant WMO expert groups such as WGNE and the Joint Working Group on Forecast Verification Research (JWGFVR).

Thank you for this suggestion and for stating the need to join up with WMO activities. We will consider this over the coming year. We are always keen to engage and align with the international community. We also recognise that our Momentum partners can contribute significant expertise in this area.

FA-R3: The Met Office should leverage its expertise and reputation in forecast verification to influence the choices of metrics used by tech companies and other AI/ML model providers to optimize and verify their forecasts.

We want to be open with private companies about standards/ metrics that they might want to use to create good forecasts and have established some interactions with the tech industry to support

this. We will take the above recommendation into account as we continue to collaborate with private companies and will be vocal and open about what good accuracy looks like on our public-facing channels, advocating high standards for all and encouraging AI based forecasts to be introduced to the market responsibly.

Hazards to Decision-Making: Met Office Warnings & Resilience Services

The National Severe Weather Warning Service (NSWWS) risk matrix approach is world's best practice for impact-based warnings. MOSAC commends the Met Office on its excellent relationship with the resilience community, and efforts to ensure that those partners are familiar with the risk matrix and comfortable with probabilistic and uncertainty concepts.

Overall, the NSWWS is very effective as measured by surveys of the public following high impact events (for example, 95% of survey respondents acted based on the warning for Storm Darragh). However, yellow warnings are problematic since they can arise from many combinations of likelihood and impact. The example given for Storm Bert highlighted the difficulty of warning when a high impact event is likely to occur somewhere within a region, but the spatial and temporal detail is unknown. Although a yellow warning was appropriate in this case due to low likelihood, some people who were severely impacted by flooding (i.e. experienced high impact) were unsatisfied with the yellow warning.

HDM-R1: Further social science work is recommended to understand how best to communicate and nuance yellow warnings for the public to prompt the desired protective actions. The "medium-impact yellow" terminology (i.e. putting the impact up front) could be one option among several to be explored.

Addressed in response to HDM-R3

HDM-R2: Evidence has repeatedly shown that providing quantitative probabilities can help people to make better decisions. The Met Office should consider how to include the numerical likelihood of an event in its warning messages, making use of social science to optimize the construction of the message.

Addressed in response to HDM-R3

HDM-R3: Given the difficulty of accurately forecasting the details of multi-hazard events, and the need to manage expectations of the public, social science should be conducted in ways to enhance public trust in the Met Office during and after highly unpredictable situations.

We are pleased that MOSAC has acknowledged the critical roles of social science and risk communication in realising the value of the NSWWS and the science which underpins it. Finding the optimum ways to communicate different risks – graphically, textually, numerically etc. - in order to inform user decision-making and to trigger the appropriate response is a key component of the Common Warnings Framework being developed by the Met Office and the UK's flood agencies. We agree that recent events suggest a more 'operational' use of social science may be required at times to help preserve the trust of the public and other stakeholders.

As a high impact event gets closer in time, the details of timing, location, severity, and impacts become more certain. Many weather services issue local short-fuse warnings based on nowcasts. The development of PLUVIA is providing improved short-fuse predictive capability for heavy rain events. Social research is needed on the "downstream" part of the warning value chain to better define the local vulnerabilities and exposure and address issues of public trust

(e.g. when the regional warning is yellow or non-existent, but the local warning is amber or red). This requires collaboration of Met Office experts with other social scientists, delivery partners from the resilience community, and government and local agencies.

HDM-Q1: MOSAC would like to know more about the Met Office's plans for whether/how to develop short-fuse impact-based warnings that account for local (potentially dynamic) vulnerabilities and compounding/antecedent conditions.

Our current priority is to build on the success of the Flood Forecasting Centre's Rapid Flood Guidance (RFG) trial in summer 2024. The Met Office science and operational capabilities developed to support the RFG have natural applications for other nowcasting scenarios and for other services. We are undertaking research with partners to improve our Surface Water Flood impact modelling capabilities, with a specific focus on better representing exposure, vulnerability and antecedent conditions. We are also making initial steps to understand the role AI/ML may play in impact modelling, in support of impact-based warnings.

The above discussion highlights the critical importance of exploiting social science to help deliver the best societal outcomes from the investment in the value chain of observations, modelling, post-processing, and forecasting science (which are concerned with producing the most accurate information). The Met Office's social scientists intimately understand the MO's business, perform highly relevant work to demonstrate and enhance the value of Met Office services (particularly for high impact weather), and serve as the interface to the broader community of UK and international social science experts.

HDM-R4: The Met Office should continue to invest in, and potentially grow, its in-house social science capability.

See response to HDM-Q1 above.

Space Weather

Space Weather was considered in depth by MOSAC for the first time at this meeting. The Space Weather Research team (SWR) is a research-to-operations and operations-to-research group that is tasked with developing improvements to the products and activities of the Met Office Space Weather Operations Centre (MOSWOC). At present the duties of MOSWOC are primarily as defined by the UK National Risk Register, but there are several other potential user communities in government and industry that may be of relevance in future. A plausible scenario is that the occurrence of a high impact event, such as a G5 level geomagnetic storm, could lead to a sudden urgent demand for increased guidance. This is a potential opportunity, and the future space weather research requirements will depend on how the Met Office wishes to position itself as new users and service providers enter the market.

The Space Weather Research team is a small group that seeks to leverage work done in the USA and in the UK academic sector, rather than developing new models and data sets on its own. The main vehicle for this research since 2019 has been the Space Weather Instrumentation, Measurement, Modelling and Risk (SWIMMR) program, which has provided funding to university partners to develop the components of a complete forecasting chain from heliosphere to the thermosphere/ionosphere. This program appears to have been highly successful, although the real test will be the integration and coupling of the components into an operational forecast chain that will take place over the next 12-18 months. Not only have the modelling components been delivered, but also an active community of collaborators has been established. There are,

however, numerous challenges associating with deeper coupling of the different parts of the chain, and with the assimilation of existing and new observation types

SW-R1 MOSAC congratulates the SWR team on the success of the SWIMMR program and recommends that this community be maintained and developed, for example through a follow-up project.

We recognise these challenges and SWIMMR follow-on support contracts have been put in place with the lead institutes for the SWIMMR modelling projects, to ensure the SWR team has ongoing support for the integration and to ensure these models can be successfully integrated by the September 2026 target date.

Now, the SWR group is rather decoupled from other research activities at the Met Office. While there are many potential areas where space weather could benefit from the expertise of other groups, resources are limited, and it is important to identify a few projects where large impact could be obtained with modest effort. As examples, MOSAC would like to call attention to two areas, where progress in SWR is slowed by lack of the necessary expertise.

SW-R2 New ground and aircraft-based instruments have been developed as part of SWIMMR but plans to deploy and operationalise these observations appear to be relatively undeveloped. The experience and knowledge of the observations research experts could be invaluable, both for the deployment of the instruments, but also for establishing pre-processing and data flows.

We note that progress is being made in this area, for example in the deployment and operationalisation of neutron monitors developed during the SWIMMR programme, but agree that this could be more widespread. We agree that this will benefit from the input from Met Office observations experts and will seek to engage with them to ensure that we can onboard further new instruments effectively. As an example, the space weather programme is now funding a post in Observations Operations to support space weather instrumentation operationalisation.

SW-R3 As the new models from SWIMMR come online, there will be numerous opportunities to increase the variety and accuracy of Space Weather forecast products. Applied Science has extensive experience in the identification of new users and co-development of suitable products, that could help to raise the profile of the Space Weather group and to guide its future priorities

We welcome this comment since we are aware of the need to develop more end-user focused products and of the relative lack of skills within our team to do so. We had already begun to consider whether we should engage with Applied Science experts and MOSAC's comments gives us further impetus and motivation to do this.

Nowcasting extreme rainfall events and their impacts

MOSAC is pleased to note that the development of a field-based precipitation analysis, nowcasting and verification framework, PLUVIA Mesoanalysis, has made good progress in producing very short-range forecasts that are closer to observations than NWP. It appears to be a good tool to assist operational forecasters in their extreme precipitation warning decisions.

NERE-R1: MOSAC recommends that ML be invested as a priority in nowcasting to improve the 6-hour forecast of mesoscale convective systems and heavy precipitation for the UK. The ability of ML methods to both learn from observational data and represent uncertainty across spatial and

temporal scales makes it a powerful method for developing new types of operationally useful nowcasts, at resolutions and lead times where alternative methods struggle.

We agree that we need to look to improve our precipitation nowcasts with some urgency. We are considering the options available to us for resourcing both essential work to replace legacy systems by PS48, and doing more innovative research, including using Machine Learning.
Recent work in the Met Office on using ML for aviation nowcasting applications has hugely increased our understanding of both the challenges and opportunities in this area. What we have learned supports our view that there are very few quick wins in using ML for nowcasting and that we need to be mindful of the balance between our ambition and the resource available to meet the hard deadlines we are facing in the coming two years. It may be beneficial to work with partners with more experience and expertise in Machine Learning, and exploiting Met Office expertise in providing high quality data sets, such as our Mesoanalysis, that can be used for training, testing and verifying ML nowcast models.

The prospects for extending PLUVIA Mesoanalysis to 3D by incorporating additional observations from MTG-IRS, MODE-S and MTG-FCI appear very promising.

NERE-R2: For nowcasting and issuing thunderstorm warnings, satellite and radar data are not always sufficient to identify and track the most active parts of damaging thunderstorms, and lightning data can be an important source of information. MOSAC encourages the consideration of lightning products as a part of the PLUVIA Cell Tracker that have received little attention, such as the MTG-LI lightning imaging mission, which is expected to provide significant benefits for nowcasting.

We agree that the inclusion of lightning information within the PLUVIA Cell Tracker will be valuable for nowcasting the hazards from thunderstorms. The inclusion of lightning data from our ground-based network, LEELA, is in our plans. This will serve to validate the predictions of the lightning nowcast from PreFLITS (implemented as part of the PLUVIA Cell Tracker) and support future prediction methods. Work is ongoing in Space Applications and Observations R&D to understand the performance of the MTG-LI, and we have no doubt that in time it will be a valuable source of lightning information for nowcasting.

MOSAC welcomes the organisation of testbeds to assess the performance of tools and nowcasts, and to facilitate interaction with users in the development of products and the introduction of new data.

Ensembles in seasonal prediction

MOSAC remains impressed by the scientific advances in seasonal forecasts. New insights of underlying mechanisms of seasonal predictability were presented. Met Office remains a world leader in seasonal forecasting with a state and science of the art seasonal forecast system. MOSAC applauds that performance indicators visualized in scorecards, tailored to seasonal forecasting, are used to objectively determine the quality of the seasonal forecast system and its updates.

MOSAC also values the deeper analysis of spread and error relationships. These are not trivial, and a nuanced view is needed with physical insights and insights of model biases of the phenomena that are predicted.

Some intriguing results were shown of changes in extreme hot conditions from uninitialized ensembles. It will be interesting to learn whether the changes in extremes follow from an extreme value distribution with a changing mean or whether feed backs can be identified that lead to departures from such a distribution.

MOSAC also considers the applications of seasonal forecasts to be of key relevance to the Met Office mission. In this respect it would welcome insight of how seasonal forecasts are used in downstream services, learn how the customers are involved and how these impacts future development of the seasonal forecast system.

The increase of computational capabilities provides opportunities for increased ensemble size, which is shown to be beneficial. MOSAC would also welcome considering alternative strategies, such as developing limited high-resolution forecasts and using AI/ML to generate ensembles. The use of AI/ML may also be of value in the postprocessing of the data.

ESP-Q1 How do the changes reported on changes in hot extremes compare to statistical expectations derived from extreme value distributions?

Thank you for this insightful question. We are indeed looking into this already and agree that it's important. At the moment we are testing a simple model based on a shift in the mean (or location parameter) according to the trend. It is not yet clear if this gives a good match but we will determine if this is the case or if additional changes to the shape of the distribution are needed due to additional feedbacks as suggested.

ESP-R1 MOSAC recommends highlighting the applications of seasonal forecasts and recommends that further development relates to the customer needs in addition to development resulting from scientific and computing advances.

We are engaging with the Department for Energy Security and Net Zero (DESNZ) as they already use our forecasts for energy resilience planning and we are investigating how they can extract more benefit from both our seasonal forecasts and the associated ensemble analysis of extreme events that we showed in the presentation. The other major users include a number of other National Met Services who take data feeds and the Copernicus Climate Change Service who also receive real time data feeds. We also have a dedicated International Sub-Seasonal to Seasonal Outreach (ISSO) team in the STEP area of Programmes (Science Technology and Engineering in Programmes), which undertakes capacity building and outreach activities on seasonal forecasting and its applications, predominantly through the WMO Regional Climate Outlook Forums (RCOFs) and associated user-engagement forums.

Momentum Model

MOSAC is pleased to see the progress discussed last year on Open Sourcing Momentum is delivering later this year and believes the choice of license (BSD-derived) and distribution platform (Github) will help foster a community of developers and users external to the Met Office.

MOSAC is pleased to see good progress with scientific performance of LFRic, and the encouraging news about the confidence in computational cost decreases over the next few months. MOSAC is also impressed with the promise shown in so short a time by SWIFT, both in terms of scientific and computational performance. Noting though that one of the drivers of LFRic was increased scalability relative to the UM, MOSAC feels that it should be a priority to revisit the testing to examine the extent to which this has been achieved.

MM-R1 To get as a complete a picture as possible of computational performance relative to the UM, MOSAC recommends revisiting scalability testing as soon as HPC resources permit.

We agree with this recommendation. It should be highlighted that assessment of the scalability of the dynamical core of LFRic (GungHo) has been explored (in collaboration with EPCC) using resources available on Archer2 (https://arxiv.org/html/2409.15859v1). This shows good scaling. However, this doesn't capture interactions with the physics schemes, is constrained to a simple meteorological setup and doesn't include I/O. These aspects are critical to providing confidence in scalability of the system, so we will endeavour to carry out further scalability tests of the more complete system as soon as HPC resources become available.

However, notwithstanding the good progress, there is still a huge amount of work to do synchronising JADA developments and the latest iterations of LFRic, optimizing performance and ultimately performing the full resolution trials with results that will enable an operational switchover.

MM-R2 To reduce the risk of slippage (and not do too much nugatory work fine tuning LFRic with the VAR system), MOSAC encourages every effort to be made to bring together JADA work and LFRic work as soon as it is practicably possible.

We agree with this assessment of risk and with the recommendation. JADA development is already using LFRic and will update to the latest science configurations of LFRic as and when they become available. For now, assessment of model forecast skill relies on the VAR system, and will use JADA as soon as practicable. The appropriate links are in place between the NG-R2O, NG-PAO and GC6 projects to ensure that plans are coordinated.

Momentum data assimilation

MOSAC congratulate the Satellite and Surface Assimilation team on significant progress in implementing JEDI and proactively anticipating potential issues. They also have an exciting program of new developments planned for the regional system.

The development of two parallel potential global data assimilation schemes (Control-Pert and EIDA) is considered by MOSAC as not an efficient strategy.

DA-R1 MOSAC encourages the Met Office senior management to decide as soon as possible about the time available in the operational schedule for data assimilation, bearing in mind key evidence on the value of initial conditions for NWP skill, and the value obtained by other centres, such as ECMWF, who can have outer loops and longer observation windows due to longer compute times.

We thank MOSAC for their acknowledgement of the progress with JEDI and we welcome this recommendation as we agree that longer compute times for data assimilation will be beneficial to the overall NWP skill as demonstrated by ECMWF. We would be happy to report back at a future

MOSAC the decision on computer resources for data assimilation and progress on the related strategy for the global data assimilation scheme.

MOSAC recognises some risky dependencies between the current lack of a stable version of LFRic and the ability to develop a model adjoint for the data assimilation system. Ongoing collaboration between DA and model development teams is commendable and strongly supported to mitigate this risk.

DA-R2 MOSAC recommends strong collaboration between the data assimilation and model team for the rest of the development cycle.

We share MOSAC's concern and welcome their recommendation. We are working hard to ensure continued effective collaboration between DA and the model development / LFRic infrastructure teams. The need for a stable version of LFRic is key for progressing the development and tuning of JADA system. Until this is achieved, regular (fortnightly) meetings take place between DA and modelling teams to capture, understand and take action around any new developments that may have knock-on impacts.

The connection between AI for National Capability and DA team expertise in observations, quality control, uncertainty, etc., and analysis training data is not currently well-defined. MOSAC would like to see a more collaborative approach to developing their training data strategy that leverages in-house expertise, while also considering the other pressures of NGMS development. MOSAC acknowledges that there are some new activities underway that promote this type of collaborative work and looks forward to hearing more about this in future meetings.

DA-R3 MOSAC recommends that the Met Office DA team collaborates on a training data strategy with the AI for National Capability team.

We thank MOSAC for this recommendation and agree that our strategy for AI/ML training data must leverage in-house expertise from across the Met Office. Currently our DA experts are fully focused on the work required to develop and implement JADA. This is complex and high priority work which requires their full attention. They will, of course, engage with the AI work as and when they have the capacity.

K-scale

MOSAC recognises the Met Office's strength in developing km-scale weather and climate models. The recent publication of RAL3 is a breakthrough, validating the advanced physical parameterisations at tropical and temperate latitudes. In addition, the development of unified physics for global and regional modelling at the kilometre scale is an essential paradigm for the future of "convection-permitting" physical modelling.

MOSAC commends the Met Office for establishing the global-regional model hierarchy using the Unified Model seamless modelling framework and model development foundations. MOSAC supports the principle of maintaining the focus on a model hierarchy approach to underpin hypothesis-driven research, rather than a more singular focus on km-scale global model development. It strongly supports the contribution to DYAMOND3 and ongoing K-scale research, with a continued focus on a model hierarchy approach.

K-R1: MOSAC urges further scientific evaluation of downscaling and cloud organisation, in line with the study already undertaken to characterise convective organisation in West Africa.

We welcome MOSAC commendation of our approach towards K-scale research, in particular on exploring a more hypothesis-driven approach underpinned by global-regional model hierarchy foundations (including need for more coupled aspects over the period ahead building on GC and REP developments).

Generation of new K-Scale data, both through the year-long hierarchy experiments for DYAMOND3 and global coupled Trailblazer testing will offer many opportunities for further scientific evaluation of both downscaling and upscaling aspects of our prediction capabilities (including in the representation of cloud organisation), and their suitability as sources of ML training data. The initiation of a UK academic project (UPSCALE) to align with Met Office K-Scale R&D is therefore particularly timely.

K-R2: Since the predictions are strongly dependent on the behaviour of the physics of the model, of the dynamics and on the interaction between the two, MOSAC wonders why the effective resolution appears so coarse and suggests that this issue be investigated in order to obtain a better resolution of scales. It also emphasises the importance of exploring the K-scale hierarchy with the LFRic-based Momentum model and its new dynamics and numerical schemes.

We agree with this encouragement to continue to assess the effective resolution across our modelling capabilities, and in particular with a view to the impact of moving to LFRic. Work is in progress to push LFRic to higher model resolutions and we are eager to explore the new scientific hierarchy that it brings. Investigation into effective resolution in the global LFRic model at lower resolutions has shown spectra are well captured and maintained, so higher resolution simulations will provide some interesting data to explore.

K-R3 MOSAC also strongly recommends machine learning emulators using K-scale simulations that could be applied to coarser resolution, in conjunction with the innovative work on hybrid AI-physics modelling.

We thank MOSAC for highlighting this. We will continue to progress work in line with this recommendation to understand the impact of K-scale simulations on data-driven ML and more hybrid approaches. We anticipate that contributions to the DYAMOND3 protocol will provide a key resource for this.

Topical presentations

- Assimilation of Mode-S observations into the Met Office global deterministic NWP model by **Elliott Warren**

This was an excellent presentation on exciting new developments in assimilation of Mode-S aircraft data. The Met Office has partnered with FlightRadar24 to collect global data from air traffic management reports. These are processed at EMADDC to provide wind and temperature observations along flight tracks and are complementary to the European data already provided by EMADDC. Experiments had been carried out to optimize the use of these data (e.g., thinning distance). Large increases in forecast skill were demonstrated, roughly equivalent to 1/3 of a typical parallel suite upgrade. The Met Office staff involved are to be congratulated on this impressive work!

- Aerosol classification from sun photometer and lidar data using machine learning by **Joelle Buxmann**

In another application of machine learning, MOSAC enjoyed hearing about the success of an aerosol classification scheme that uses sun photometer data from AERONET and a k-nearest neighbour technique. Compared to a validation dataset of lidar retrievals of aerosol type, the sun photometer classifications were 92% accurate. These results are very encouraging for using near-real time AERONET data to nowcast air pollution, smoke, dust. MOSAC was happy to see that the Met Office plans to apply this approach to nowcast volcanic ash, a serious hazard for aircraft.

- A new ensemble post-processing technique for physically-consistent spot forecasts by **Alice Lake**

Alice Lake showed how to add value to deterministic site forecasts using ensemble information. She explained how the mode of a multivariate distribution identifies the most probable prediction for a multi-risk hazard, but loses physical consistency. This can be restored by choosing a representative ensemble member near to the mode. which can be interpreted physically in the same way as a deterministic forecast. This work provides an excellent example of how ensembles can benefit customers that cannot make direct use of probabilistic predictions.

- Identifying ensemble perturbations that tip the stratosphere into a sudden warming by **Chris Kent**

The MOSAC was intrigued by the presentation on sources and consequences of sudden stratospheric warmings. Using a creative workflow with ensembles and numerical model perturbation techniques, distinct large-scale conditions were found in the North Atlantic preceding sudden stratospheric warmings, also hinting towards a bifurcation of the stratospheric circulation. The study provides important scientific insights on the drivers and predictability of sudden stratospheric warmings and nicely leverages the ensemble generation capabilities of the Met Office that allows conducting such studies.

- The SWIFT Transport Scheme for LFRic by **Thomas Bendall and James Kent**

This very interesting talk presented the new monotonic flux-form semi-lagrangian (SL) transport scheme developed for LFRic. A major advantage of this new scheme is that it holds for large Courant numbers. Therefore, it leads to a significant reduction in the cost of horizontal transport, which is very good news for LFRic. It was developed for tracers, with the monotonicity, but has been extended to all variables (without being monotonic or positive-definite), which is another major advance. The associated publication has also shown that SWIFT performs better than the COMIC scheme on the varying wind or varying density test cases, and is comparable to PPM, which is a great achievement. The local conservation of moisture and monotonicity have allowed the global balance of moisture to be improved, in addition to the fact that mixing ratios are now

used for moisture variables, which also improves conservation. It will be very interesting to measure its impact on convective events, such as during the WesCon campaign, when there were too many showers, probably linked to excessive vertical velocities: the scheme could potentially help to reduce this effect. It would also be interesting to measure improved accuracy and see if the effective resolution is finer. This new scheme is a big step forward for LFRic and could be used in other operational SL models.